Successfully Integrating NASA Data Into an On-going Public Health Study and Linking NASA Environmental Data with a National Public Health Cohort Study to Enhance Public Health Decision Making

UAB

Leslie McClure (PI)

Shia Kent

George Howard

Matthew Loop

USRA

Mohammad Al-Hamdan (co-PI)

William Crosson (Co-I)

Maury Estes(Co-I)

Sue Estes (Co-I)

Gina Wade

Sarah Hemmings

NASA

Dale Quattrochi (Co-I)

CDC

Sigrid Economou

Mark Puckett

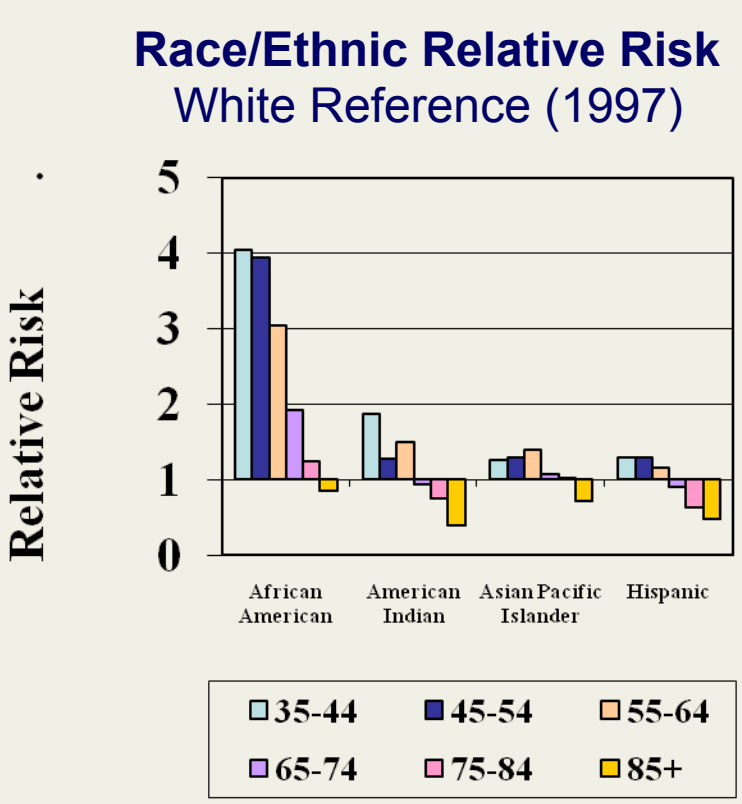
Goals and Objectives

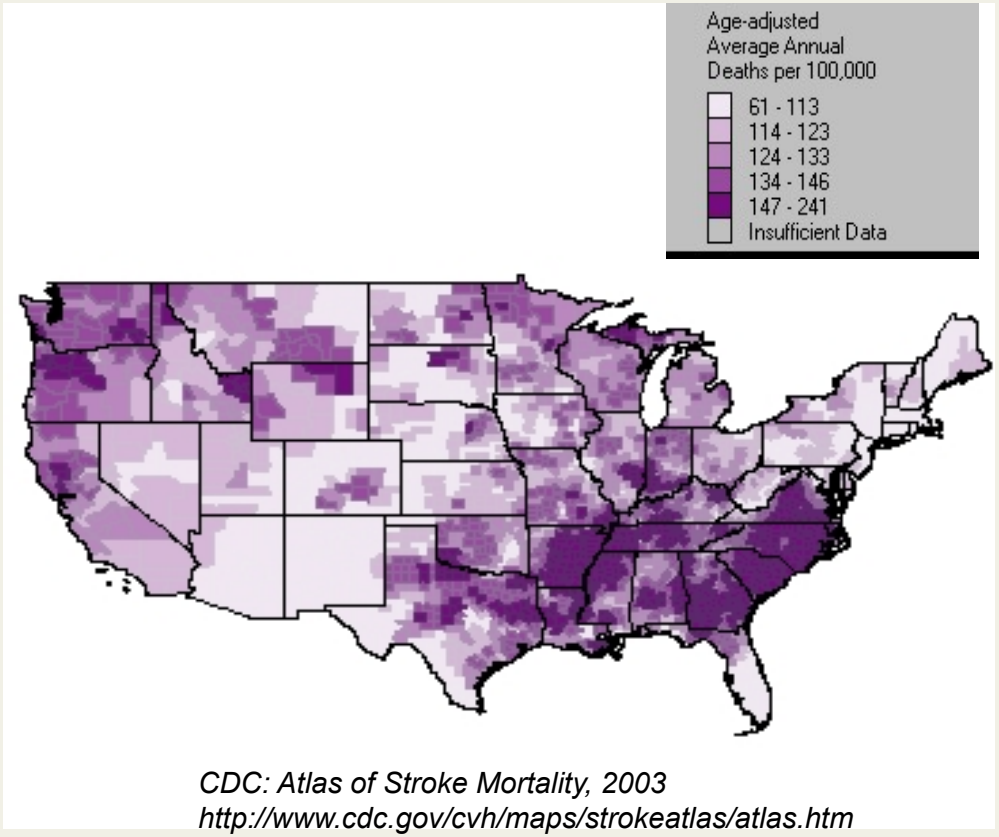
- **This project has dual goals in decision-making activities**
  - Providing information to decision makers about associations between environmental exposures and health conditions in a large national cohort study
  - Enriching the CDC Wide-ranging Online Data for Epidemiologic Research (WONDER) system by integrating environmental exposure data
- **Develop daily high-quality spatial data sets of environmental variables for the conterminous U.S. for the years 2003-2008 utilizing NASA data (Objective 1)**
  - Fine Particulates ($PM_{2.5}$) (NASA MODIS and EPA AQS)
  - Land Surface Temperature (NASA MODIS)
  - Solar Insolation and Heat-related Products (Reanalysis Data)
- **Link these environmental variables with health data from a national cohort study and examine environmental health relationships (Objective 2)**
  - Cognitive Function
  - Hypertension, diminished kidney function, hypertension, and inflammation
- **Make the environmental datasets available to public health professionals, researchers and the general public via the CDC WONDER system (Objective 3)**

Why REGARDS?

(REasons for Geographic And Racial Differences in Stroke)

Racial & Regional Disparities

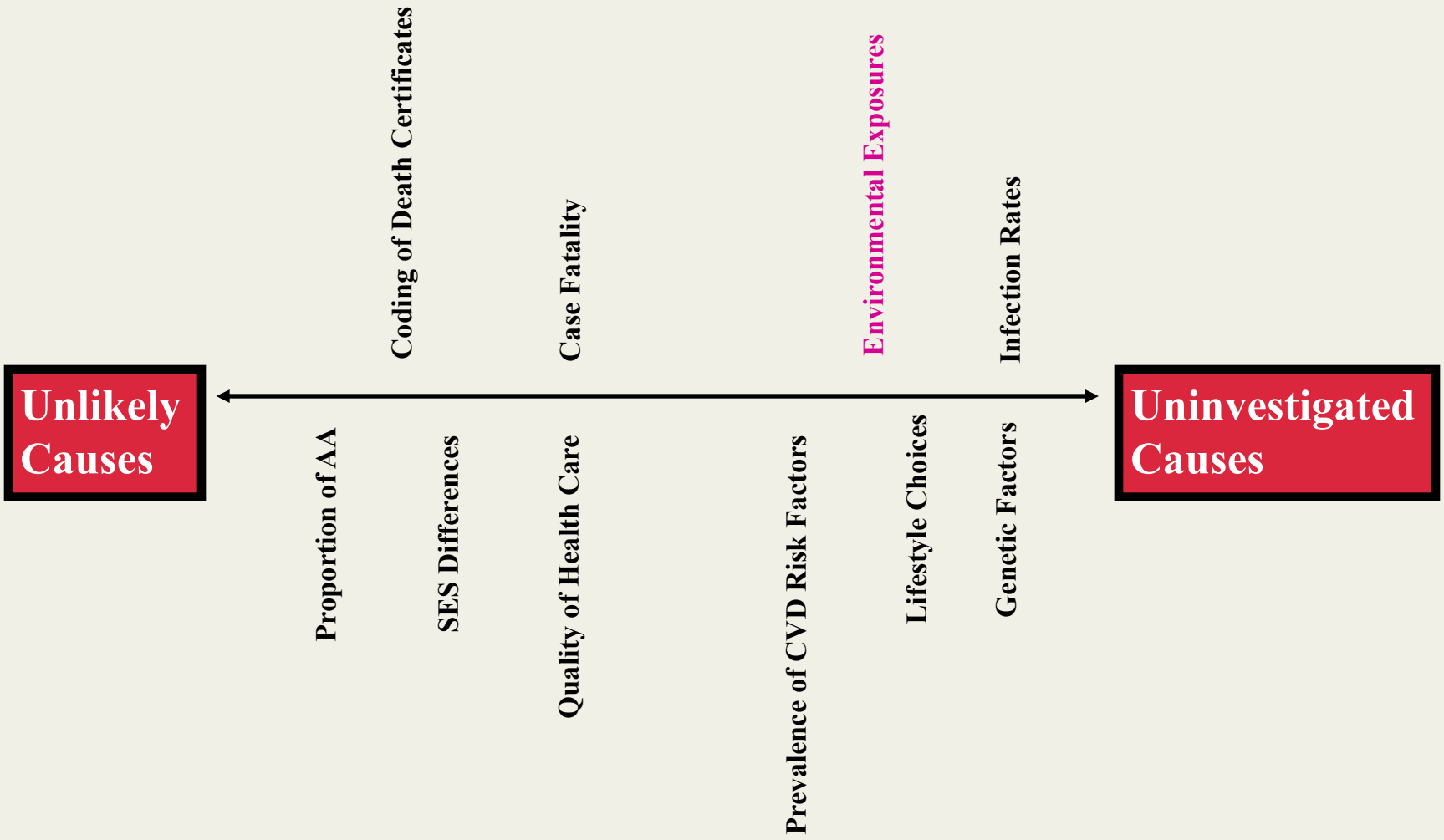




CDC: Atlas of Stroke Mortality, 2003
http://www.cdc.gov/cvh/maps/strokeatlas/atlas.htm

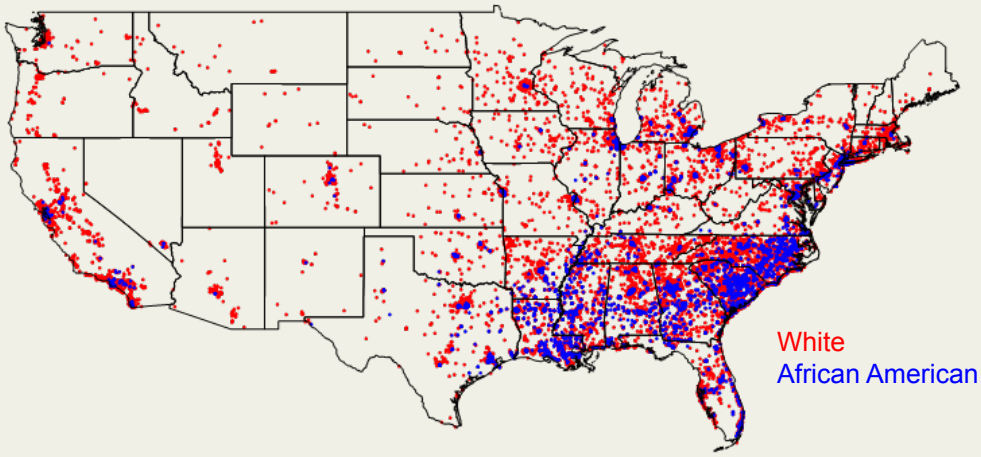
Why REGARDS?

Regional Disparity



REasons for Geographic And Racial Differences in Stroke (REGARDS) Study Population

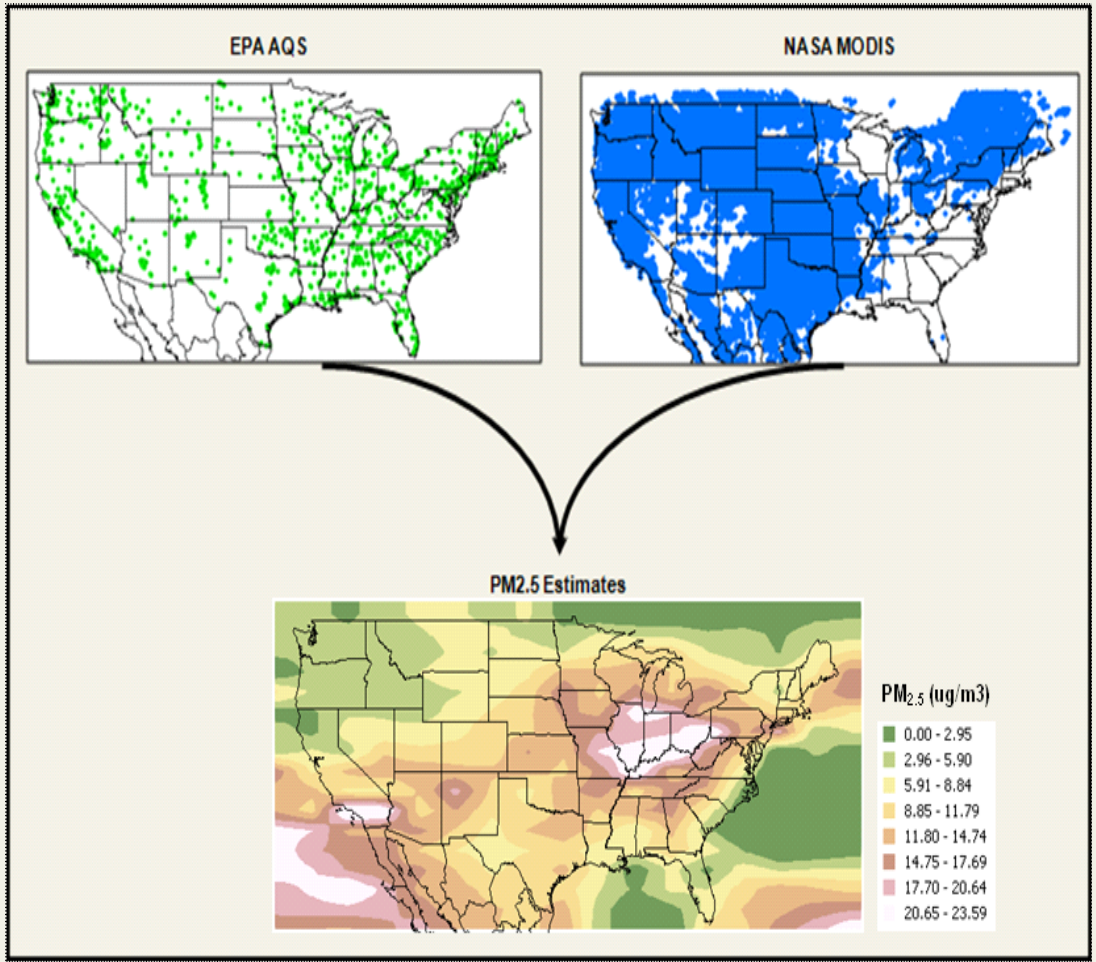
- Longitudinal population-based cohort of over 30,000 volunteers age 45 and older
- Racial representation
  - 42% African American
  - 58% white
- Sex representation
  - 45% male
  - 55% female
- Geographic representation
  - 21% from the buckle of the stroke belt
  - 35% from the stroke belt
  - 44% from the rest of the contiguous US
- Successfully transferred from UAB to NASA/MSFC
  - BAA as per HIPPA Regulations
  - Data Encryption



National Environmental Datasets (Objective 1)

Fine Particulate Matter ($PM_{2.5}$)

- Estimated ground-level $PM_{2.5}$ from MODIS AOD using published regression equations per EPA region per season (Zhang et al., 2009)
- Combined with EPA $PM_{2.5}$ data from the AQS for 2003-2008
- Modified and ran MSFC Surfacing Algorithm (Al-Hamdan et al., 2009, 2012) to produce continuous spatial surfaces of daily $PM_{2.5}$ for the contiguous US for 2003-2008



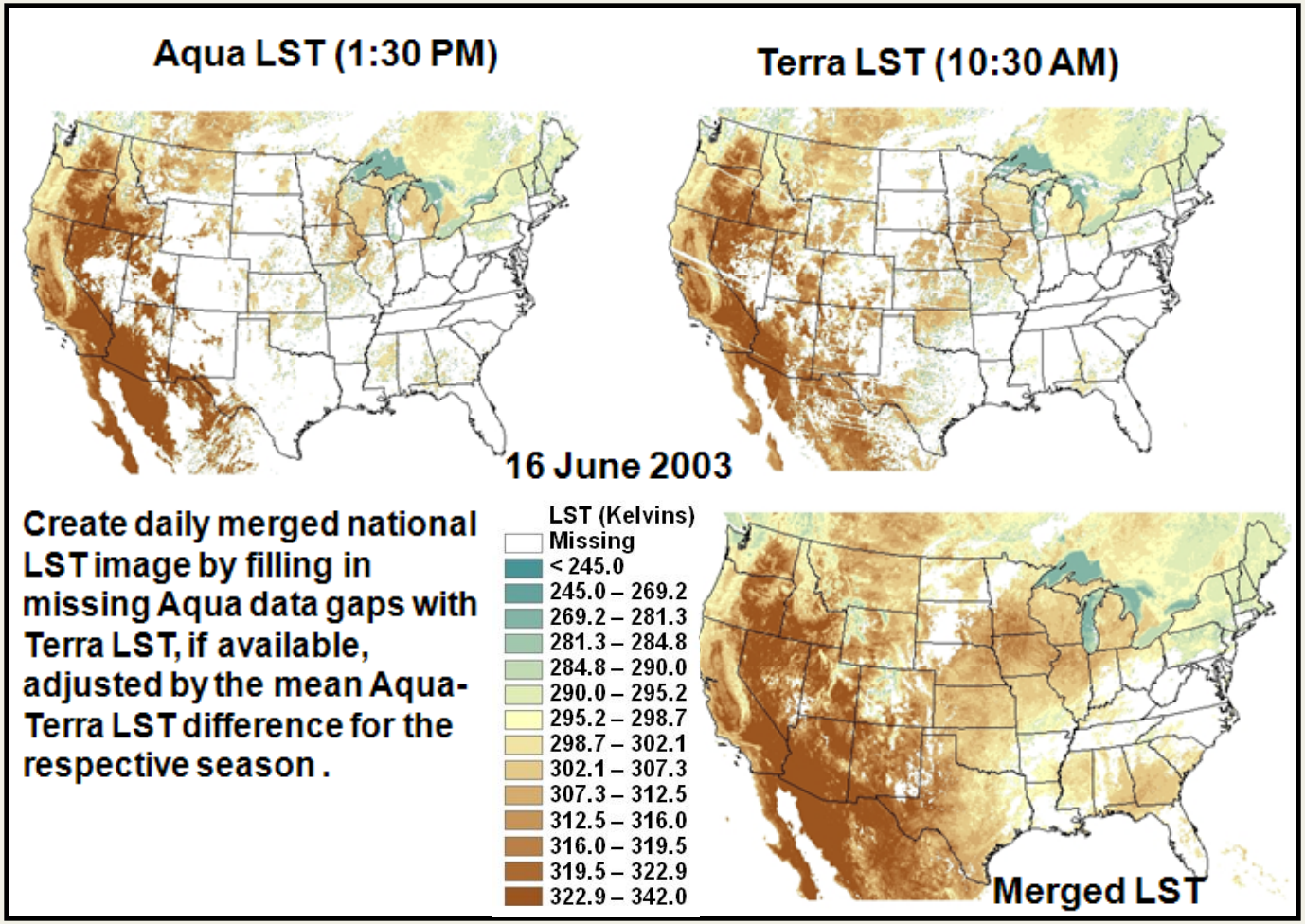
$PM_{2.5}$ on July 14, 2003

(10 km spatial resolution)

(*Al-Hamdan et al., Geocarto International, 2012*)

Land Surface Temperature (LST)

- Aqua and Terra daytime & nighttime data for 2003-2008 were processed
- Aqua-Terra differences were computed by season for 2003-2008
- Aqua data gaps were filled with Terra-adjusted LST (if available) by mean seasonal difference
- National merged Aqua-Terra daily LST dataset were generated for 2003-2008 for day & night (*Crosson et al., RSE, 2012*)

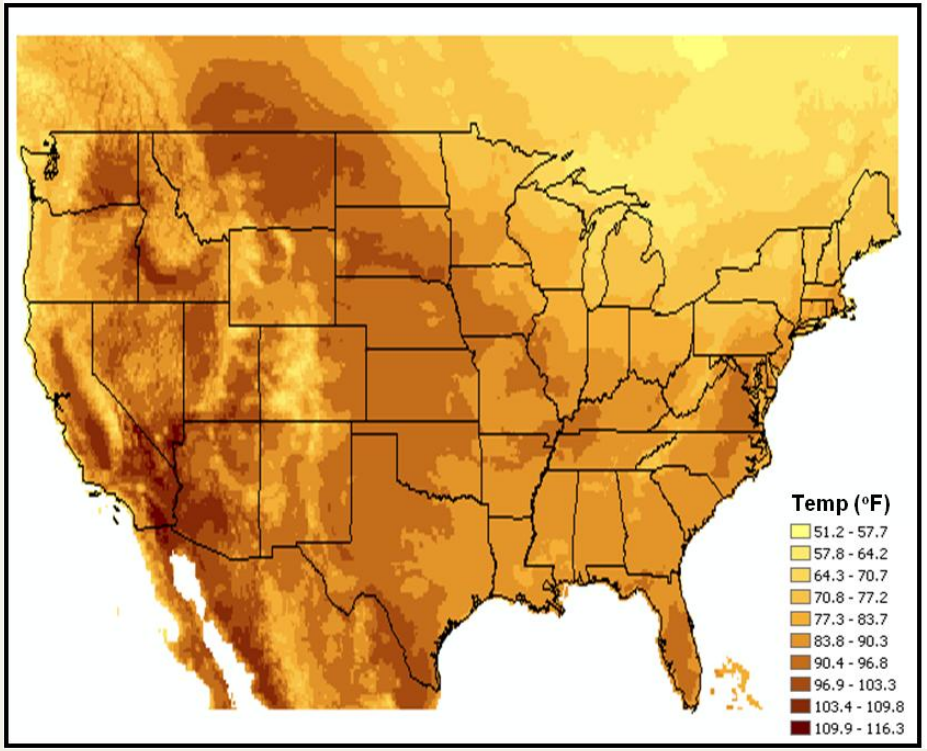


(1 km spatial resolution)

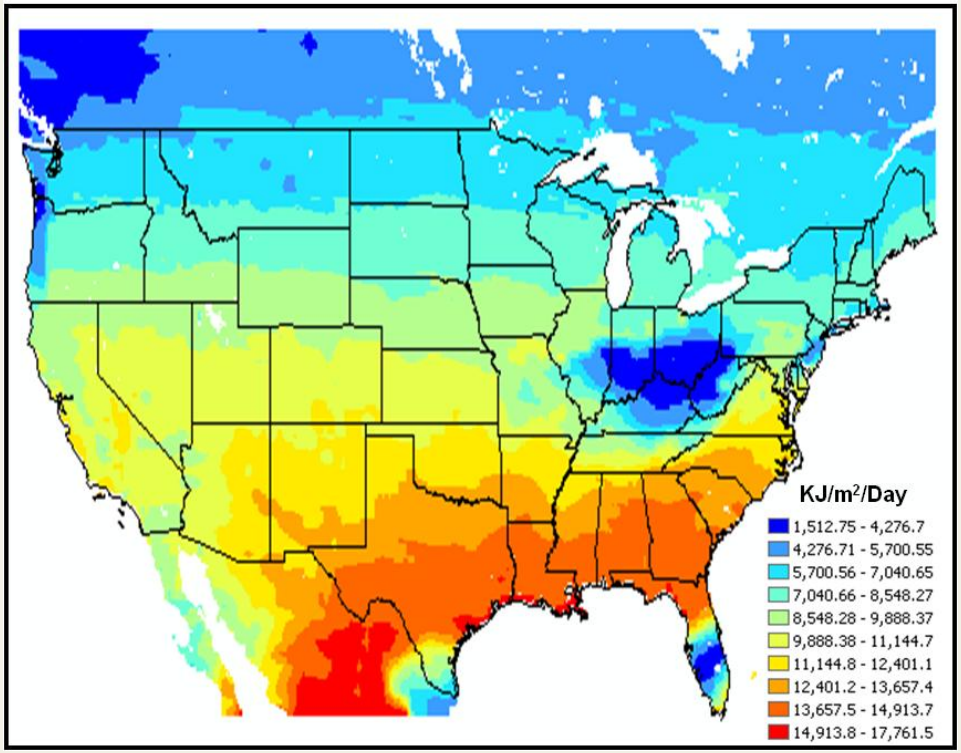
Heat and Solar Insolation

- NLDAS hourly forcing data (air temperature, solar radiation, specific humidity, atmospheric pressure) for the 2003-2008 period were processed
- Daily statistics of Maximum Air Temperature, Minimum Air Temperature, Maximum Heat Index, and Total Solar Insolation were computed for 2003-2008

NLDAS Max Air Temperature on July 15, 2008



NLDAS Solar Insolation on January 1, 2008



(12 km spatial resolution)

Final Report

1. Produce daily gridded estimates of $PM_{2.5}$ for the conterminous US for the years 2003-2008 from MODIS Aqua data
2. Produce daily gridded solar insolation (SI) maps for the conterminous US during the same period using data from the NARR (originally proposed) (Used NLDAS)
3. Produce daily gridded and surface temperature (LST) maps over the conterminous US during the same period using data from MODIS
4. Link the estimates of $PM_{2.5}$, SI and LST with data from the more than 30,000 participants from the REGARDS study.

All datasets have been produced and linked.

Environmental Health Data Linkage and Biostatistical Analyses

(Objective 2)

Data Linkage for Biostatistical Analyses

- Link in a GIS the estimates of the $PM_{2.5}$, Solar Insolation, and Air Temperature with health data from all participants in the REGARDS study on the individual level at the geographic coordinates of their residences
- Sort the environmental data by participant ID, and merge in with the corresponding health data from the REGARDS database
- Determine whether exposures to these environmental risk factors are related to cognitive decline and other health outcomes such as hypertension, inflammation, and stroke

Participant ID	Lat	Lon	Day1 Solar Insolation ($KJ/m^2/Day$)	Day2 Solar Insolation ($KJ/m^2/Day$)	Day3 Solar Insolation ($KJ/m^2/Day$)	.....	Day365 Solar Insolation ($KJ/m^2/Day$)
1	99.045	-87.105	7950	8941	8945		7850
2	99.055	-89.036	7401	8501	8412		7501
3	99.065	-86.212	8001	7015	8251		8401
. . . . .							
30200	99.075	-87.855	15650	11402	15650		10750

Simulated example of the linked data set consisting of participant ID and the associated NLDAS solar insolation

Final Report

Determine whether exposure to $PM_{2.5}$ or SI is related to the rate of cognitive decline among participants in the REGARDS study, independent of other known risk factors for cognitive decline

- SI, temperature and cognition analyses have been completed and manuscript published (Kent et al., *International Journal of Biometeorology*, 2013)

- $PM_{2.5}$ and cognition analyses completed and manuscript published (Loop, et al. *PLOS ONE* , 2013)

Solar Insolation: Cognition

- Analyses first assessed the most appropriate exposure for both SI and temperature period: 15 yr, 10 yr, 5 yr, 2 yr or 1 yr
  - Found 1 year period prior to baseline to have the strongest association with cognitive decline
  - However, no clear temporal trends emerged
- Next, assessed whether the association between each of SI and temperature and cognitive decline was significant after inclusion of known risk factors for cognitive decline

Solar Insolation: Cognition

- Found that even after multivariable adjustment, there is an association between SI and cognitive decline, but that this relationship differs depending on the temperature (p for interaction=0.0011)

Association between SI
(below median vs. above median)
and decline, by temperature tertile

	OR (95% CI)
1st Tertile of Temp	1.26 (0.94, 1.68)
2nd Tertile of Temp	1.30 (1.06, 1.58)
3rd Tertile of Temp	1.95 (1.29, 2.96)

$PM_{2.5}$: Cognition

- Assessed association between $PM_{2.5}$ and incident cognitive impairment (SIS)
- Used logistic regression to determine whether the likelihood of impairment was impacted by $PM_{2.5}$
  - Assessed this after adjusting for known factors associated with $PM_{2.5}$, as well as known factors associated with cognitive impairment

Odds Ratio (OR) and 95% CI for the effect of a 10 ug/m^3 increase on $PM_{2.5}$ on the odds of incident cognitive impairment

Model	Covariates	OR (95% CI)
M1	Length of FU, temperature, season, incident stroke	1.26 (0.97, 1.64)
M2	+M1, age, race, region, education, income	1.02 (0.76, 1.37)
M3	+M2, smoking, alcohol use, exercise, BMI	0.97 (0.72, 1.31)
M4	+M3, depressive symptoms, dyslipidemia, diabetes, hypertension	0.98 (0.72, 1.34)

Final Report

Examine the relationship between the estimated $PM_{2.5}$ and SI and other health-related conditions among REGARDS participants, including diminished kidney function, hypercholesterolemia, hypertension, and inflammation (CRP)

- Analysis of SI, temperature and stroke has been completed and manuscript published (Kent et al., *Annals of Neurology*, 2013)
- Association between SI and secondary outcomes (measures of cholesterol, hypertension, dyslipidemia, inflammation) has been examined
  - Manuscript drafted, and we are in the process of trying to publish it
- On-going analyses are assessing the association between PM2.5 and stroke

Solar Insolation: Secondary Outcomes

- To assess the impact of SI on stroke risk factors, we did a split sample analysis
  - Randomly divided the cohort in half
    - Exploratory analyses on one half, confirmatory on the 2nd
- Outcomes assessed included:
  - Continuous: HDL, LDL, total cholesterol, CRP, SBP
    - Examined using linear regression
  - Dichotomous: CKD, hypertension, dyslipidemia
    - Examined using logistic regressoin
- Only SBP, LDL and HDL were significantly different by levels of exposure in univariate models
- After multivariable modeling, and confirmatory analysis, only HDL was significantly associated with SI
  - Higher sunlight exposure associated with lower HDL levels

Solar Insolation: Stroke

- Analyses followed the same methods as the cognitive decline data
  - Found 1 year period prior to baseline to have the strongest association with stroke
  - However, no clear temporal trends emerged
- Next, assessed whether the association between each of SI and temperature and stroke was significant after inclusion of known risk factors for stroke

Solar Insolation: Stroke

- Found that those above the median SI exposure were less likely to have stroke than those below the median, independent of known stroke risk factors and temperature
- Those exposed to SI below the median were 1.61 times more likely to have a stroke than those exposed to SI levels above the median (95% CI: 1.15, 2.26)
- Found that there was a J-shaped association between maximum temperature and stroke, again after multivariable adjustment

	OR (95% CI)
1st Quartile of Temp	1.41 (0.99, 2.03)
2nd Quartile of Temp	REF
3rd Quartile of Temp	1.69 (1.17, 2.46)
4th Quartile of Temp	1.91 (1.27, 2.91)

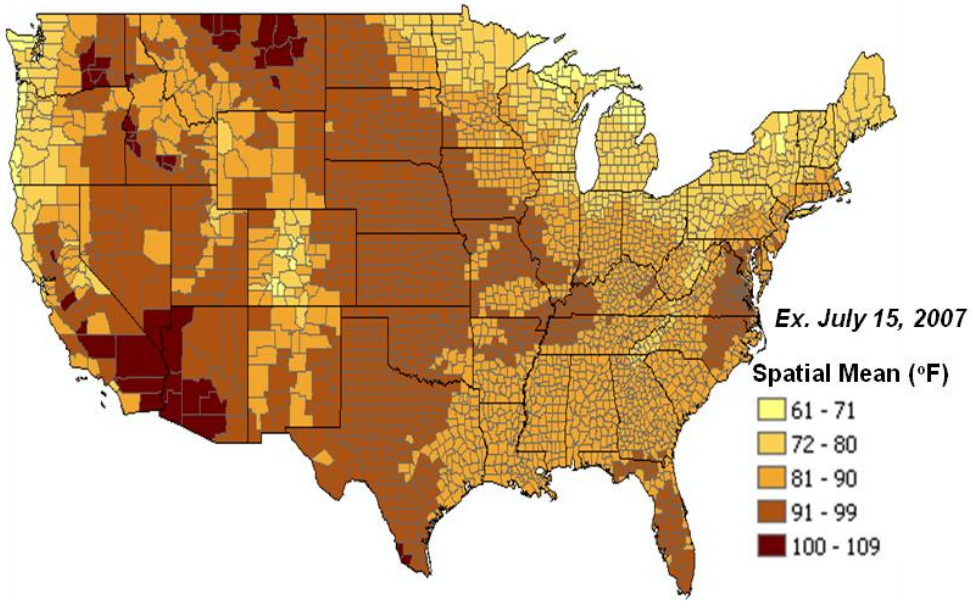
Data Dissemination via CDC WONDER (Objective 3)

Data Dissemination via CDC WONDER

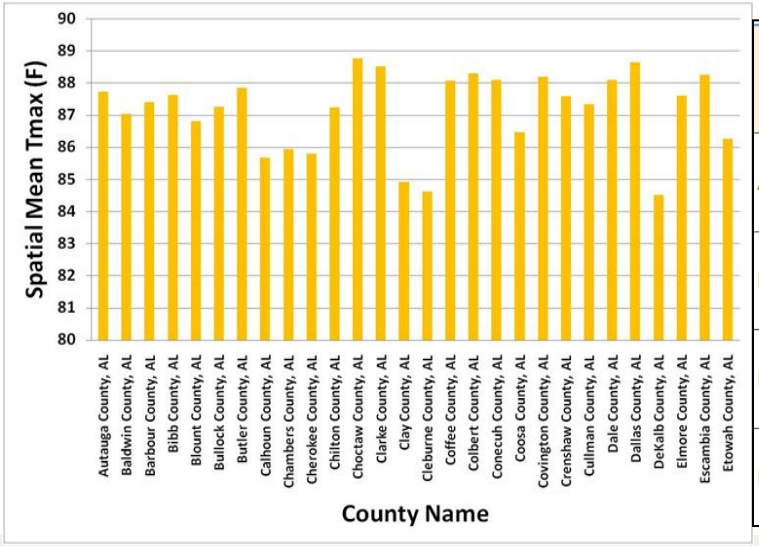
Tabular Grid-level Daily Data

Grid Cell ID	County, State	FIPS	Day1 Tmax (°F)	Day2 Tmax (°F)	Day3 Tmax (°F)	.....	Day365 Tmax (°F)
1	Kern, CA	06029	71	74	66		70
2	Kern, CA	06029	70	72	67		69
3	Kern, CA	06029	72	73	66		72
. . . . .							
103936	Aroostook, ME	23003	35	31	32		34

Examples of County-level Spatial and Temporal Statistics (Map and Chart) as provided by CDC-WONDER real-time data queries



- Environmental exposure datasets will be made available to public health professionals, researchers and the general public via WONDER, where they can be aggregated to the county-level or higher as per users' need
- Users are able to spatially and temporally query datasets and create county- and higher-level maps and downloadable statistical tables and charts of data across the *contiguous* U.S.
- Enabling easy linkage of the environmental exposure data with other health data available via CDC WONDER



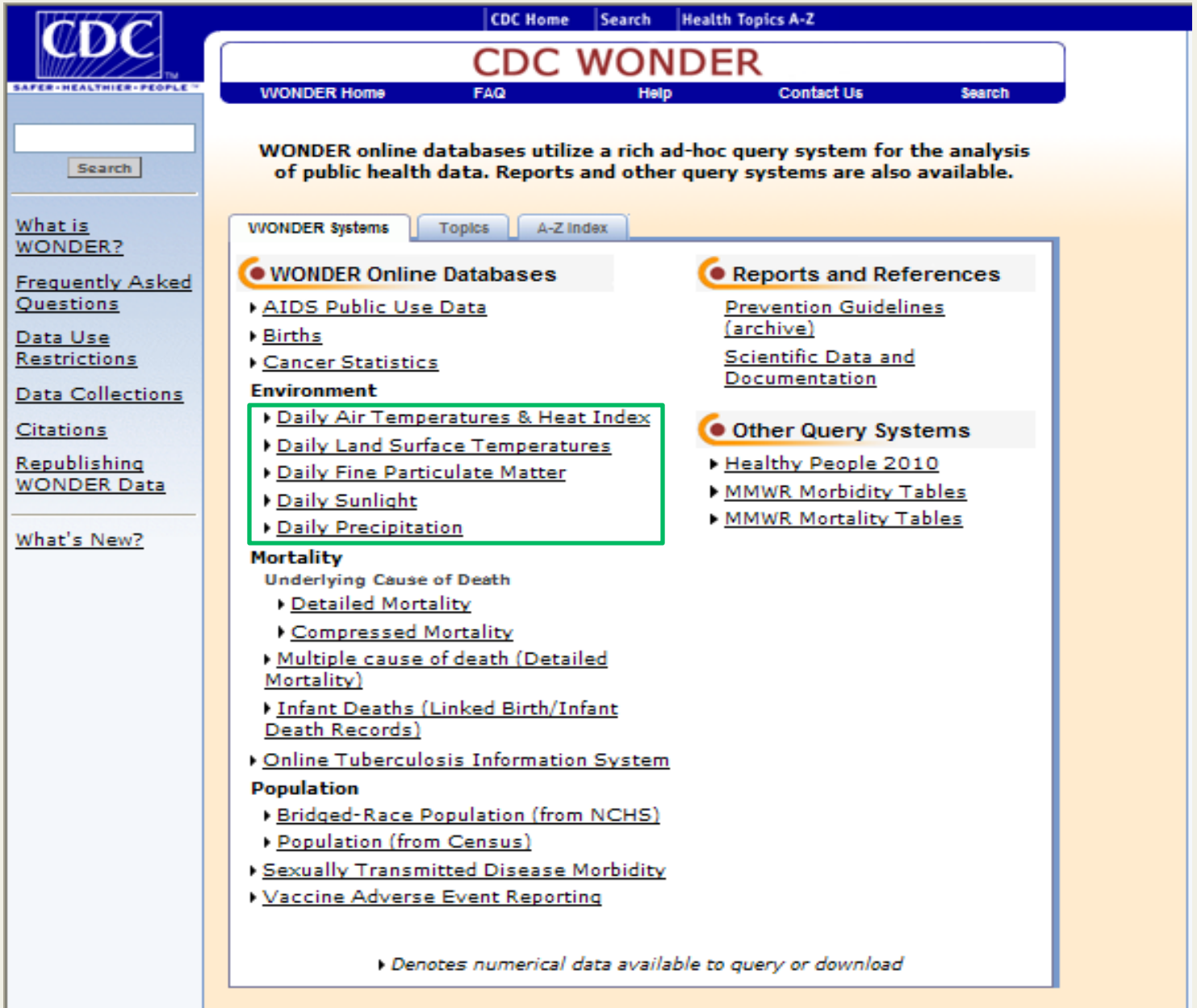
County	Avg Daily Max Air Temperature(F) # of Observations Range Standard Deviation
Autauga County, AL (01001)	87.85 11 (87.20 to 88.40) 0.43
Baldwin County, AL (01003)	85.82 26 (84.30 to 87.20) 0.61
Barbour County, AL (01005)	86.04 14 (85.50 to 86.60) 0.37
Bibb County, AL (01007)	86.92 9 (86.40 to 87.50) 0.31

Final Report

- All proposed daily national environmental datasets **have been integrated into CDC-WONDER, and are live on the website**
- Additional dates were added at the CDC and end-users' request
  - $PM_{2.5}$ (2003-2011)
  - LST (2003-2008)
  - Maximum and minimum air temperature, maximum heat index, and solar insolation (1979-2011)
- Precipitation was added at the CDC and end-users' request

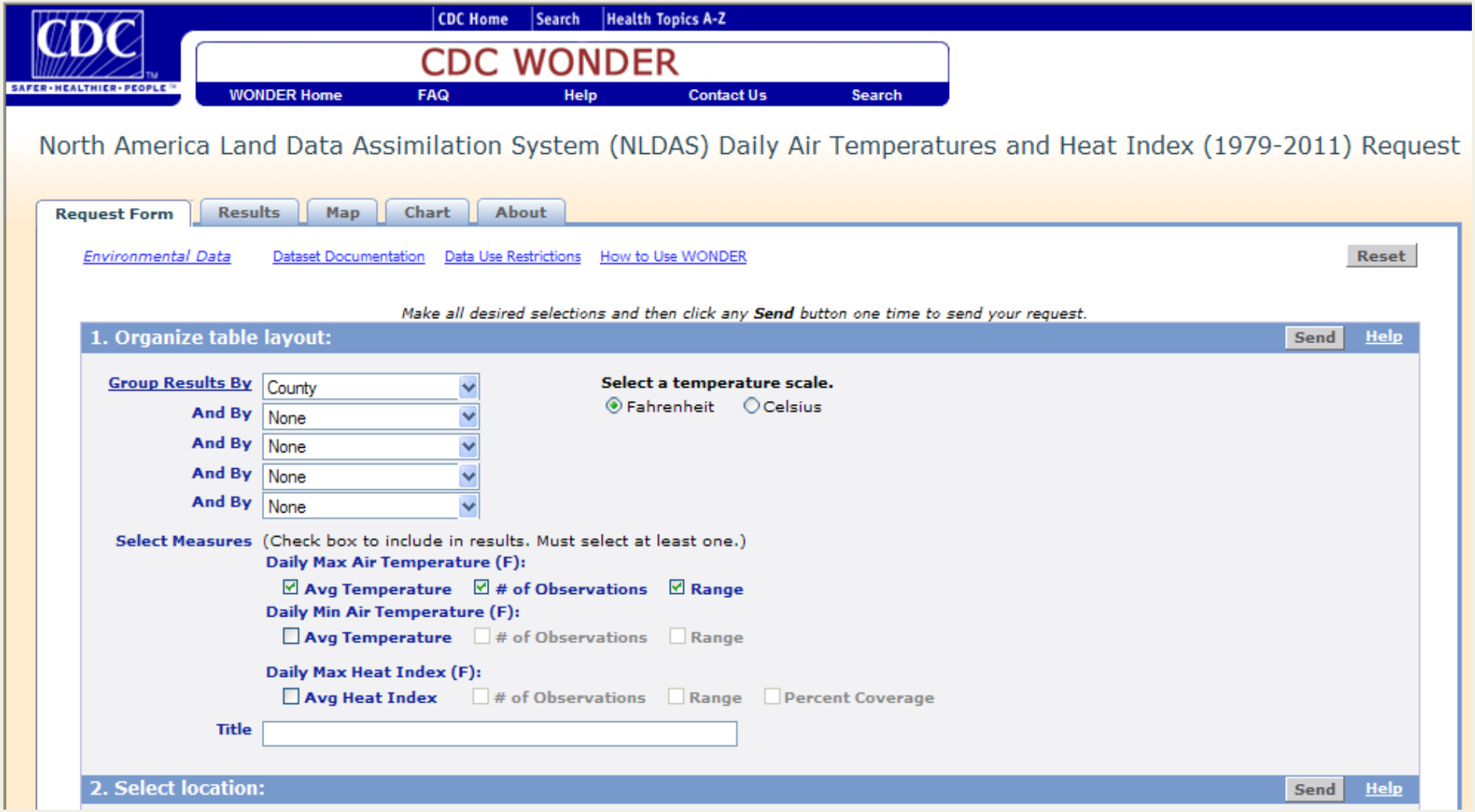
CDC WONDER Main Web Page

http://wonder.cdc.gov/



NLDAS-derived Heat-related Products on CDC WONDER

Now Available at http://wonder.cdc.gov/nasa-nldas.html



CDC WONDER Tabular Results

North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index (1979-2011)
Results

Request Form | Results | Map | Chart | About

Environmental Data Dataset Documentation Help for Results Printing Tips Help with Exports

Export Reset

Quick Options More Options

Top Notes Citation Query Criteria

County	Avg Daily Max Air Temperature (F) # of Observations Range
Autauga County, AL (01001)	87.85 11 (87.20 to 88.40)
Baldwin County, AL (01003)	85.82 26 (84.30 to 87.20)
Barbour County, AL (01005)	86.04 14 (85.50 to 86.60)
Bibb County, AL (01007)	86.92 9 (86.40 to 87.50)
Blount County, AL (01009)	84.20 10 (83.60 to 84.90)
Bullock County, AL (01011)	86.57 10 (86.10 to 87.30)

CDC WONDER Map Results

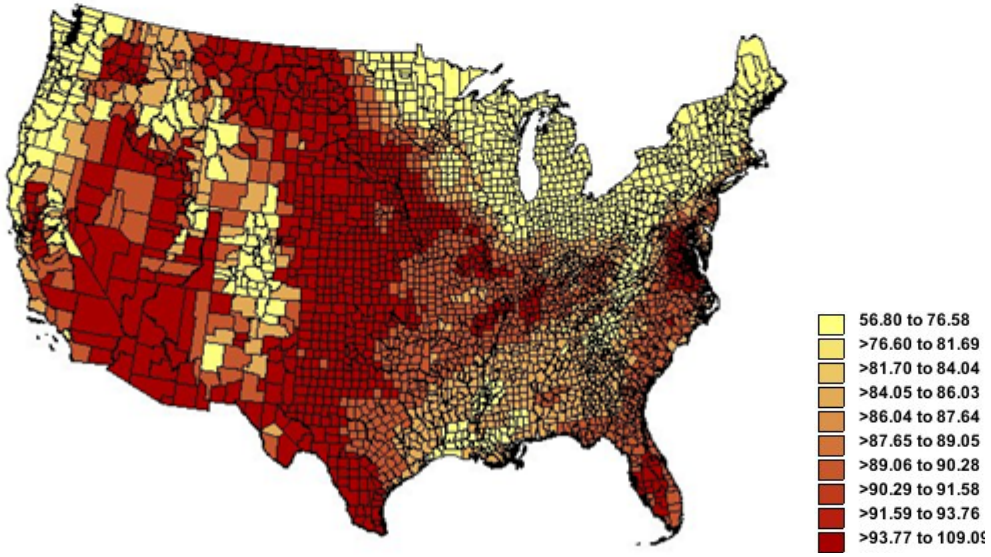
Request Form | Results | Map | Chart | About

Dataset Documentation | Help for Maps | Printing Tips

Reset

Top | Map Options | Notes | Citation | Query Criteria

Avg Daily Max Air Temperature(F) for The United States



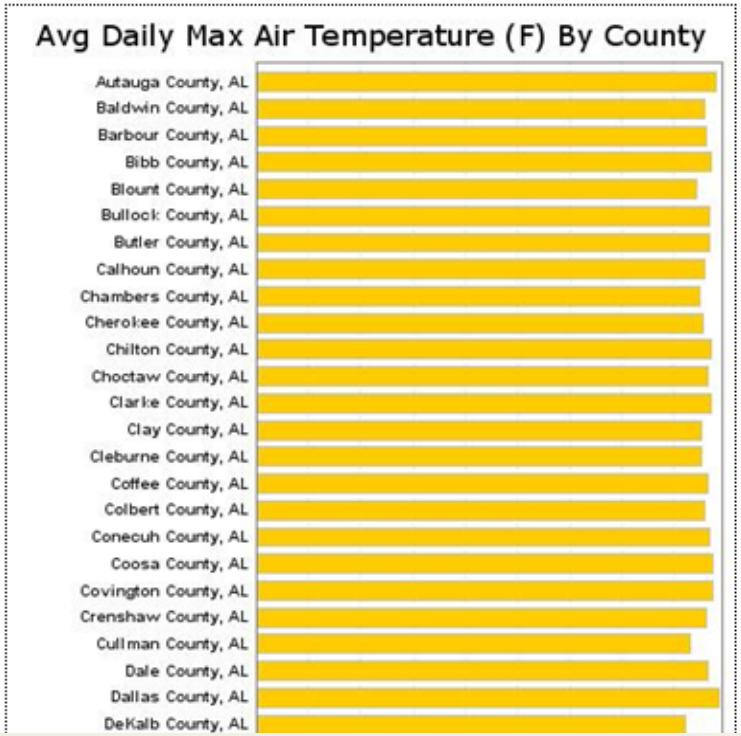
CDC WONDER Chart Results

North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index (1979-2011) Charts

Request Form | Results | Map | Chart | About

Environmental Data Dataset Documentation Help for Charts Printing Tips Reset

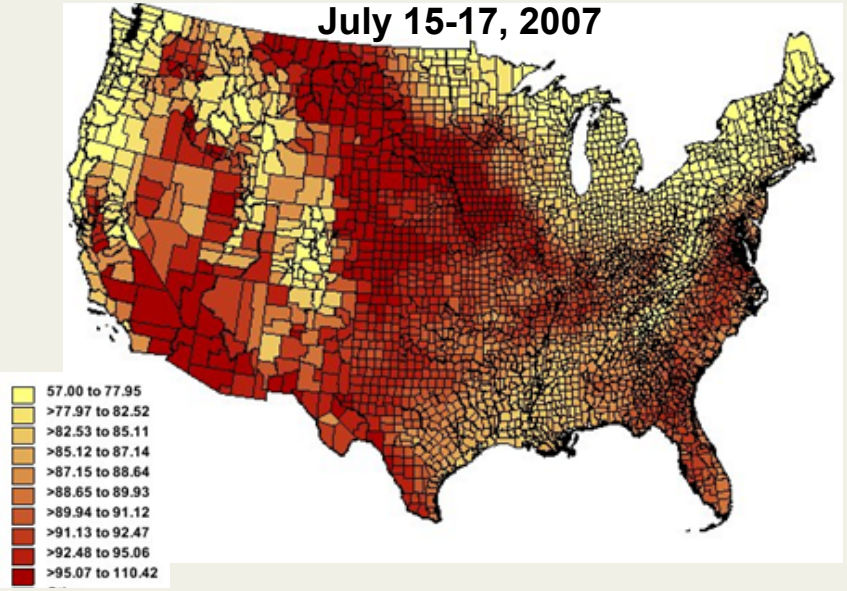
Top Chart Options Notes Citation Query Criteria



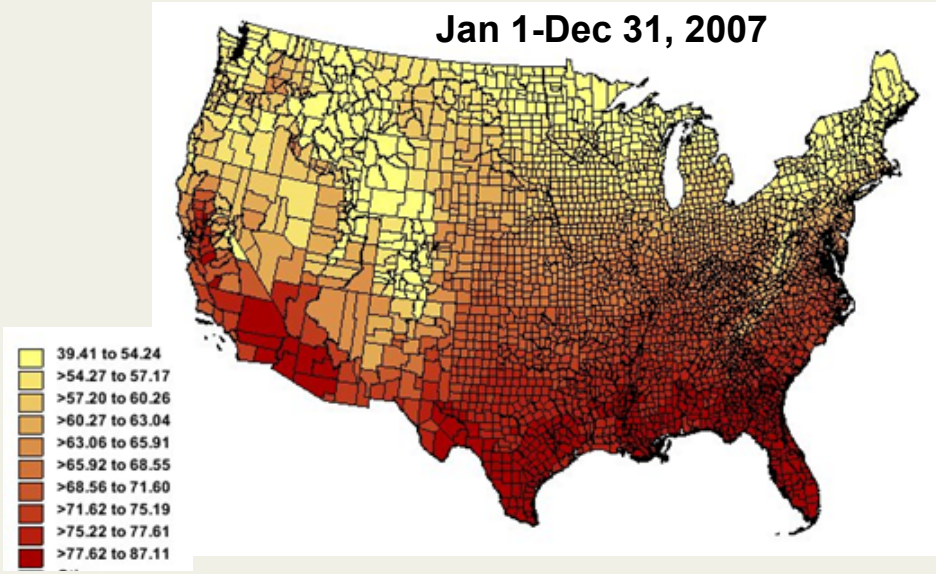
CDC WONDER Spatial/Temporal Aggregation

Avg Daily Max Air Temperature(F) for The United States

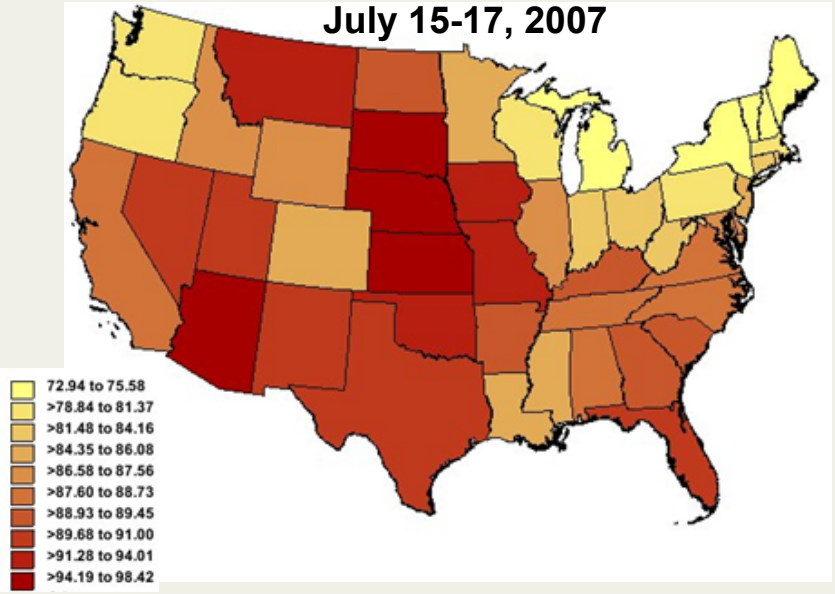
July 15-17, 2007



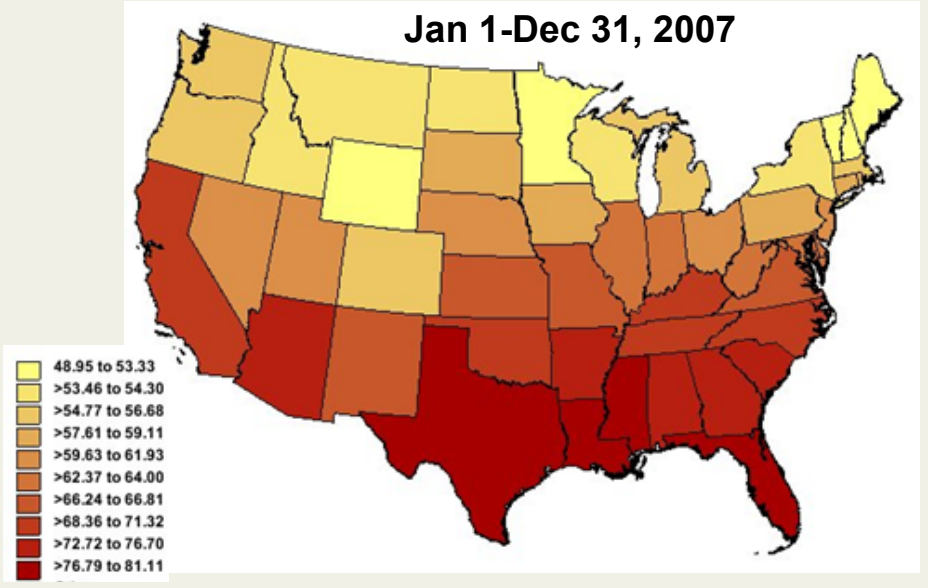
Jan 1-Dec 31, 2007



July 15-17, 2007



Jan 1-Dec 31, 2007



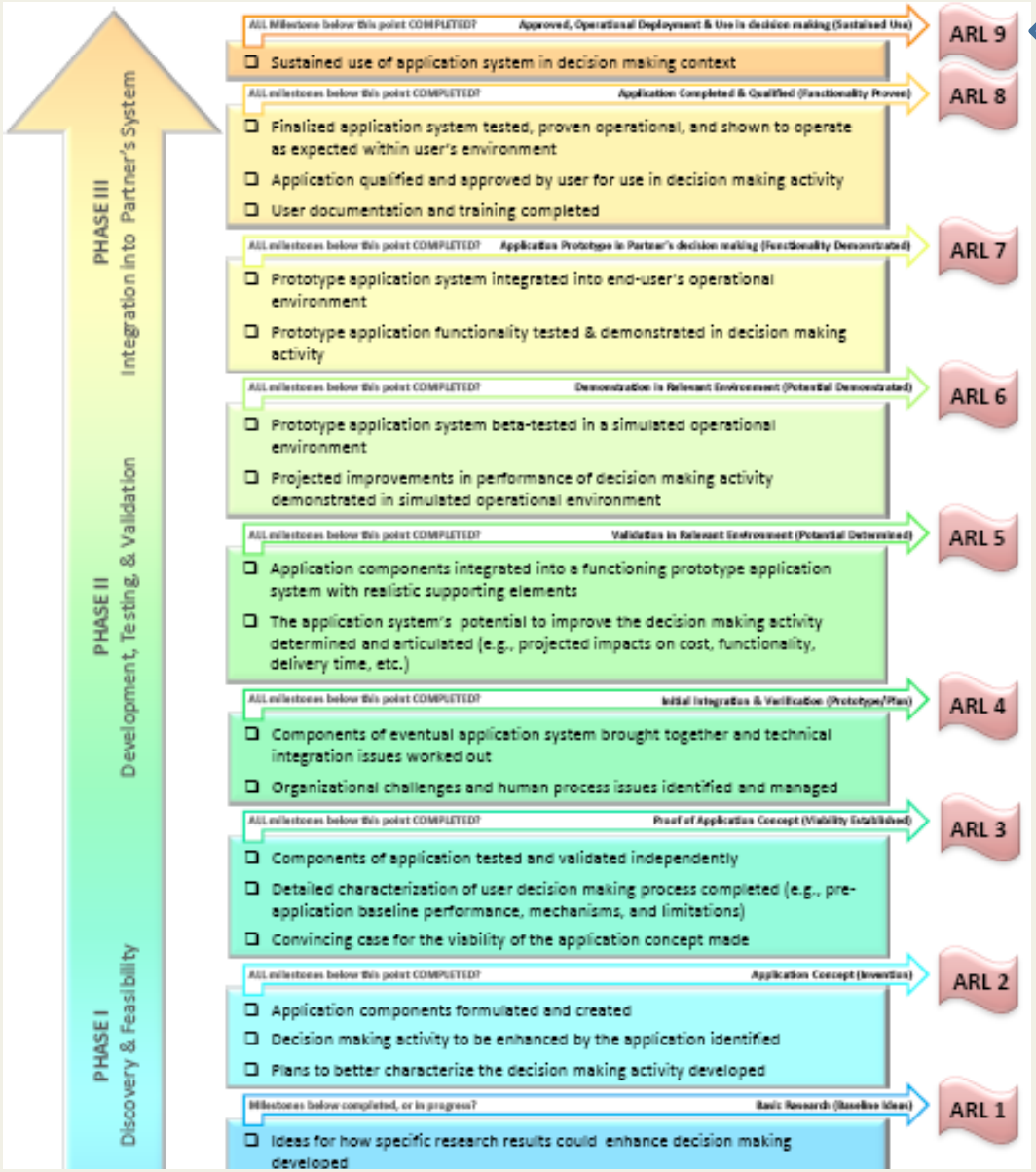
Schedule

Task		Year 1				Year 2				Year 3			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1a	Production of LST data set												
1b	Production of solar insolation data set												
1c	Production of $PM_{2.5}$ data set												
2a	Linkage of LST and cognitive decline data sets												
2b	Linkage of insolation and cognitive decline data sets												
2c	Linkage of $PM_{2.5}$ and cognitive decline data sets												
3a	Analysis of insolation and cognitive decline												
3b	Analysis of $PM_{2.5}$ and cognitive decline												
4a	Analysis of insolation and secondary outcomes												
4b	Analysis of $PM_{2.5}$ and secondary outcomes												
5	Transition to end-users through CDC WONDER												
6	Final research report												

Remaining Work

- Completion of the analyses of $PM_{2.5}$ and stroke are pending additional data
- Other aspects of cognition have now been measured in REGARDS
  - It would be of interest to determine whether both SI and $PM_{2.5}$ have associations with particular cognitive domains

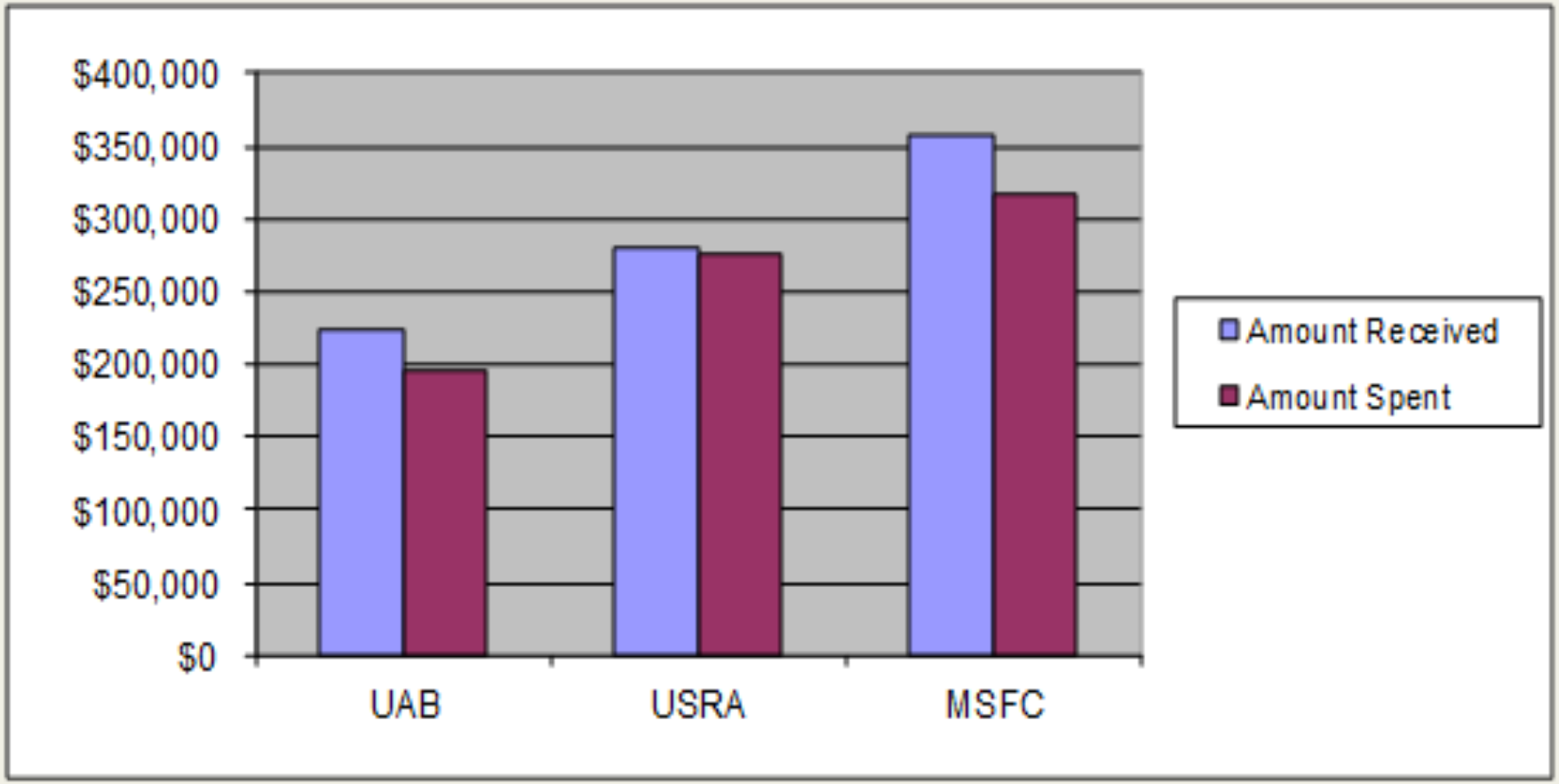
Application Readiness Level (ARL)



Financial Activity

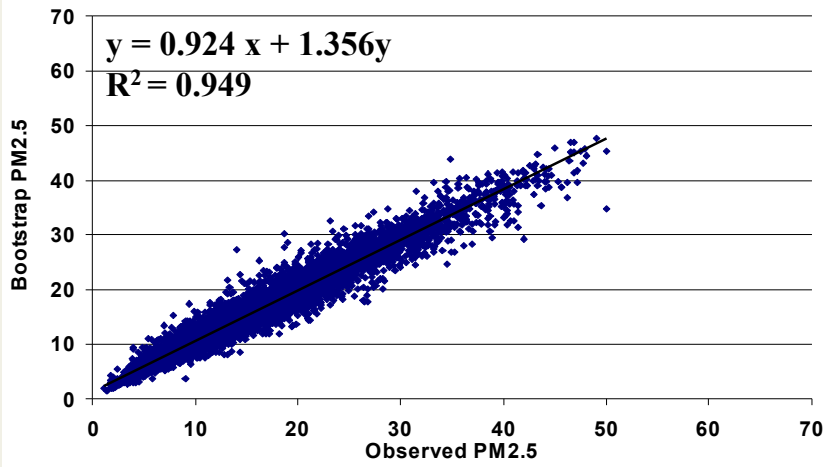
- Through the end of August, we have spent 92% of the dollars allocated for this project
  - 25% by UAB
  - 35% by USRA
  - 40% by MSFC
- In the process of transferring $10,000 to USRA to help with processing of humidity data

Financial Activity



Performance Measures: $PM_{2.5}$ Data

- **Cross-Validation: a.k.a. 'bootstrapping' or 'omit-one' analysis**
- **Objective: Estimate errors associated with daily spatial surfaces**
- **Procedure:**
  1. **Omitting one observation, create surface using N-1 observations**
  2. **Compare value of surface at location of omitted observation with the observed value**
  3. **Repeat for all observations**



Surfacing Technique and Data Source	Improvement
Bspline: AQS only vs. Merged AQS/MODIS	16 %
IDW: AQS only vs. Merged AQS/MODIS	40 %

(*Al-Hamdan et al., JAWMA, 2009*)

Performance Measures: CDC WONDER

- Through 2012, there were 79,914 hits to the environmental databases
- Through August 2013, there have been 68,835 hits to the environmental data
- It appears that the usage is increasing; however, some of the data were added in 2012, and other updated in 2012
  - Over time, we will be able to track if usage is changing

Performance Measures: CDC WONDER

Usage of Environmental Dataset on CDC WONDER

Database	Calendar 2012	Through 8/27/2013
Daily Air Temperature/Heat Index (1979-2011)	37,851	17,381
Daily $PM_{2.5}$ (2003-2011)	16,185	16,187
Daily LST from MODIS (2003-2008)	9,072	12,190
Daily Sunlight from NLDAS (1979-2011)	8,038	10,975
Daily Precipitation from NLDAS (1979-2011)*	8,768	12,102

*Prepared and provided through a different project

Endusers' Feedback

We're very excited about the feedback we've received:

- User feedback to improve the system
- Interest from Google Public Data Explorer
- Interest from Temboo Library (App developer)

Example of CDC WONDER Users Feedback

"*First, I wanted to say how very useful this dataset is. I have begun using the temperature data in my own research (heat waves and human health), and I'm planning on also using precipitation and PM values in the future.* ***Having these estimates aggregated at the county level is enormously useful in pairing up environmental and health datasets***." (Dr. Georgiana Anderson, Johns Hopkins School of Public Health, Sep. 12, 2013)

Journal Publications

- Al-Hamdan, MZ; Crosson, WL; Economou, SA; Estes Jr., MG; Estes, SM; Hemmings, SN; Kent, ST; Puckett, M; Quattrochi, DA; Rickman, DL; Wade, GM; McClure, LA. 2102. "Environmental public health applications using remotely sensed data" *Geocarto International* (in press, published on-line).
- Crosson, W.L., M.Z. Al-Hamdan, S.N. Hemmings and G.M. Wade, 2012. A daily merged MODIS Aqua-Terra land surface temperature data set for the conterminous United States. *Rem. Sens. Environ.*, 119, 315-324. doi:10.1016/j.rse.2011.12.019.
- Loop, MS; Kent, ST; Al-Hamdan, MZ; Crosson, WL; Estes, SM; Estes Jr, MG; Quattrochi, DA; Hemmings, SM; Wadley, VG; McClure, LA. "Fine particulate matter and incident cognitive decline in the REGARD study." PLOS ONE (in press).
- Kent, ST; Kabagambe, EK; Wadley, VG; Howard, VJ; Crosson, WL; Al-Hamdan, MZ; Judd, SE; Peace, F; McClure, LA. "The relationship between long-term sunlight radiation and cognitive decline in the REGARDS cohort study." *The International Journal of Biometeorology* (in press).
- Kent, ST; McClure, LA; Judd, SE; Howard, VJ; Crosson, WL; Al-Hamdan, MZ; Wadley, VG; Peace, F; Kabagambe, EK (2012). "Short and long-term sunlight radiation and stroke incidence." *Annals of Neurology* 73(1): 32-7.
- Kent, ST; Howard, G; Crosson, WL; Prineas, RJ; McClure, LA (2011). "The association of remotely sensed outdoor temperature with blood pressure levels in REGARDS: a cross-sectional study of a large, national cohort of African-American and white participants." *Environmental Health* 10(1) 7. PMCID: PMC3171279

Conference Presentations

Al-Hamdan, M; Crosson, W; Economou, S; Estes, M.; Estes, S.; Hemmings, S.; Kent, S.; Puckett, M.; Quattrochi, D.; Wade, G.; McClure, LA. "Public Health Applications of Remotely-Sensed Environmental Datasets for the Conterminous United States." International Society for Photogrammetry and Remote Sensing 2nd Symposium on Advances in Geospatial Technologies for Health/IMGA-GSA 5th International Conference on Medical Geology. August 24-30, 2013. Arlington, VA.

McClure, LA; Loop, MS; Al-Hamdan, M; Crosson, W; Kissela, BM; Kleindorfer, DO. "Fine Particulate Matter ($PM_{2.5}$) and the Risk of Stroke in the REGARDS Cohort" *International Stroke Conference* February, 2013. Honolulu, HA.

Al-Hamdan, M.; Crosson, W.; Economou, S.; Estes, M.; Estes, S.; Hemmings, S.; Kent, S.; Puckett, M.; Quattrochi, D.; Wade, G.; McClure, LA. "Using Environmental Remotely Sensed Data for National Public Health Applications." *American Geophysical Union*. December 4-7, 2012. San Francisco, CA.

Al-Hamdan, M.; Crosson, W.; Economou, S.; Estes, M.; Estes, S.; Hemmings, S.; Kent, S.; Puckett, M.; Quattrochi, D.; Wade, G.; McClure, LA. "Using NASA Environmental Data to Enhance Public Health Decision Making". *CDC Science Symposium on Climate and Health*. September 12-13, 2012. Atlanta, GA.

Al-Hamdan, M.; Crosson, W.; Economou, S.; Estes, M.; Estes, S.; Hemmings; S.; Kent, S.; Quattrochi, D.; Wade, G.; McClure, LA. "Using NASA Remotely Sensed Environmental Data in a National Public Health Study". *Urban and Regional Information Systems Association (URISA) GIS in Public Health Conference*. June 27-30, 2011. Atlanta, GA.

Al-Hamdan, M.; Crosson, W.; Economou, S.; Estes, M.; Estes, S.; Hemmings, S.; Kent, S.; Puckett, M.; Quattrochi, D.; Wade, G.; McClure, LA. "Linking NASA Environmental Data with a National Public Health Cohort Study and a CDC on-line System to Enhance Public Health Decision Making." *American Meteorological Society Annual Meeting*. January 22-26, 2012. New Orleans, LA.

Kent ST, Kabagambe EK, Wadley VG, Howard VJ, Crosson WL, Al-Hamdan MZ, Judd SE, Howard G, Peace F, McClure LA . (2012) "The Relationship Between Long-Term Sunlight Radiation and Cognitive Decline in the REGARDS Cohort Study". International Society for Environmental Epidemiology Conference. August 26-30, 2012. Charleston, SC.

Kent, ST; McClure, LA; Howard, VJ; Crosson, WL; Al-Hamdan, MZ; Wadley, VG; Judd, SE; Peace, F; Kabagambe, EK (May, 2011). UAB Center for Cardiovascular Biology Retreat, Birmingham, AL, "The use of ground and satellite data to determine the relationship between long and short-term sunlight exposure with stroke incidence in the Reasons for Geographic And Racial Differences in Stroke (REGARDS) Study." (poster- awarded1st prize in the population sciences division).

Conference Presentations (Cont.)

Kent, ST; et al (October, 2011). “Use of ground and satellite data to determine the relationship between long-term sunlight exposure and stroke incidence in the REasons for Geographic And Racial Differences in Stroke (REGARDS) Study” APHA National Meeting, Washington, DC.

Kent, ST; McClure, LA; Howard, VJ; Crosson, WL; Al-Hamdan, MZ; Wadley, VG; Judd, SE; Peace, F; Kabagambe, EK. “The use of ground and satellite data to determine the relationship between long and short-term sunlight exposure with stroke incidence in the REasons for Geographic And Racial Differences in Stroke (REGARDS) Study.” *International Stroke Conference*, February 1-3, 2012. New Orleans, LA.

Additional Notes

- Outside the scope of this project, we continue to collaborate on other projects
  - Submitted an R01 application to NIH to examine additional environmental risk factors in REGARDS
  - Collaborating with Dr. Henry Wang at UAB to prepare an R01 application for February 2014 submission, examining $PM_{2.5}$ as a risk factor for sepsis infections

Summary

- Development of national daily products of $PM_{2.5}$ (2003-2011), LST (2003-2008), maximum and minimum air temperature, maximum heat index, and solar insolation (1979-2011)
- Linkages of national environmental data with health data from the REGARDS national cohort study for environmental health correlation studies
- Dissemination of these environmental datasets to public health professionals, researchers and the general public via the CDC WONDER online system **http://wonder.cdc.gov/**
- Providing a significant addition to CDC WONDER, allowing public health researchers and policy makers to better include environmental exposure data in the context of other health data available in CDC WONDER online system
- Substantially expanding public access to these NASA environmental datasets, making their use by a wide range of decision makers more feasible

Thanks!

Questions?